UNIVERSITY OF PENNSYLVANIA

THE BEST FITTING FREQUENCY FUNCTION FOR URBAN'S LIFTED WEIGHT RESULTS

HAROLD MARSHALL LUFKIN

A THESIS

IN PSYCHOLOGY

PRESENTED TO THE FACULTY OF THE GRADUATE SCHOOL IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

PHILADELPHIA
1928

THE BEST FITTING FREQUENCY FUNCTION FOR URBAN'S LIFTED WEIGHT RESULTS

A THESIS

IN PSYCHOLOGY

PRESENTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF PENNSYLVANIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

HAROLD MARSHALL LUFKIN

Offprint from

The American Journal of Psychology, Vol. XL, No. 1

PHILADELPHIA

1928

R. UNIV.
BIBLIOTHEEK
LEIDEN

THE BEST FITTING FREQUENCY FUNCTION FOR URBAN'S LIFTED-WEIGHT RESULTS

By H. M. LUFKIN, University of Pennsylvania

Urban,[1] in assuming the $\Phi(\gamma)$-hypothesis, gives as the conditions which must be fulfilled (1) that the probabilities for a 'heavier' judgment must be expressed by a function f (x), which is asymptotic both to the x-axis and to the line y=1, or that

$$\lim_{x' = -\infty} f(x) = 0 \quad \text{and} \quad \lim_{x' = \infty} f(x) = 1$$

(for judgments of 'lighter' the limits are reversed); and (2) it is evident that this function must be expressed by means of an integral

$$y = \int_{-\infty}^{\infty} F(x)dx$$

which must exist and have a finite value greater than zero for all values of x from $-\infty$ to $+\infty$, and that the probability for any particular comparison stimulus shall be given by

$$p = \frac{1}{M} \int_{-\infty}^{x} F(x)dx.$$

There are many functions which would fulfill these conditions, but the one selected is

$$F(x) = e^{-G(x)}.$$

This function G(x) is expressed in the form

$$G(x) = h^2(x - a)^2$$

where x is the variable stimulus, a the limen, and h is a measure of precision. In this form

$$p = \frac{1}{M} \int_{-\infty}^{x} e^{-h^2(x-a)^2} dx$$

where $M = \int_{-\infty}^{\infty} e^{-h^2(x-a)^2} dx$. Urban shows that

$$M = \frac{\sqrt{\pi}}{h}.$$

*Accepted for publication April 7, 1927.

[1] F. M. Urban, Die psychophysischen Massmethoden als Grundlagen empirischer Messungen, *Arch. f. d. ges. Psychol.*, 15, 1909, 261-355; 16, 1910, 168-227.

This hypothesis has as its basic assumption that the probabilities arrange themselves in the form of the integrated probability function. If such a condition is true it will simplify the computations necessary to determine the constants which are desired. This can be done by the Method of Least Squares, but it must be remembered that this method is based on the assumption that the *arithmetical mean* is the most probable value that can be assumed for a series of observations.[2] From a series of observation equations, each of which is to be properly weighted,[3] the constants a and h can be obtained. The difference between the limens for observations of 'heavier' and 'lighter' is termed the 'interval of uncertainty,' in which neither the 'heavier' or 'lighter' judgments have a probability as great as 0.5. Urban[4] has found a rather distressing condition here, due to the fact that his probable errors for this interval are surprisingly large. This has led to the belief that the fundamental assumption of the $\Phi(\gamma)$-hypothesis is not valid and also that Urban's material, instead of being homogeneous, is heterogeneous. It is the purpose of this paper to determine these two points.

Since the publication of Urban's first paper a vast amount of work has been done in the field of statistical analysis, particularly in that branch devoted to 'frequency functions.' Karl Pearson[5] has developed his types of frequency functions, which are quite well known but as yet little used in psychometric work. There is also another type of function which is now coming into prominence, *i.e.* the Gram-Charlier. It is our purpose to examine Urban's original data and see if either of these functions are fundamentally better than the assumption of the 'probability integral.'

Since the frequencies of a judgment of 'heavier' or 'lighter' for any given comparison stimulus are in the form of an integrated function it is necessary to take the successive judgments for each value of the comparison stimulus. This gives us our values so that they can be handled as a frequency distribution. We may either work on the basis of the number of judgments or on the percentage of judgments. Whichever method is used, the results will be the same. Here the work has been done on the number of judgments, in order to simplify to some degree the work of curve fitting.

The use of the normal or Gaussian function necessitates only the computation of the first and second moments, that is,

---

[2]L. D. Weld, *Theory of Errors and Least Squares*, 1921, 65-69.
[3]Urban, *loc. cit.*; W. Brown and G. H. Thompson, *Essentials of Mental Measurement*, 1921, 65-69.
[4]Urban, *loc. cit.*
[5]Karl Pearson, Skew variation in homogeneous material, *Phil. Trans. Roy. Soc.*, A, 186, 1895, 343 ff.; 197, 1901, 443-459; 216, 1916, 429-457.

the Arithmetic Mean and the Standard Deviation or σ. In the case of the Pearson and Gram-Charlier functions it is necessary to compute in addition to these the third and fourth moments about the mean, which is done in a manner analogous to the computation of the first and second moments. This raises at once a serious objection to the use of those two types of functions. Fisher[6] quoting Charlier says that it is generally useless to compute moments of a higher order than the second when the number of observations is less than 1000. Thiel gives the following rules:

For the first and second moments rely exclusively on the observed data.

For moments between the second and sixth rely partly upon theory and partly upon observations. These objections could

TABLE I

Constants for the Pearson Curves

O	Type	β_1	β_2	k_2	Range	Skewness	Mode
I H	J_I	1.052±.168	3.488±.220	−0.67±.06	82−106	−1.56	108.0
II H	I_I	0.452±.123	3.461±.342	−0.87±∞	36−112	−0.39	102.1
III H	I_I	0.735±.125	3.433±.260	−0.01±.16	76−109	−0.73	104.2
IV H	III(VI)	0.354±.164	3.566±.515	+4.13±∞	−116	−0.30	101.2
VI H	I_I	0.730±.185	3.618±.409	−0.07±.58	64−110	−0.60	103.7
VII H	I_I	0.711±.250	3.985±.717	−3.87±∞	−0−111	−0.44	102.9
I L	I_I	0.109±.038	2.656±.116	−0.08±.03	82−109	+0.25	90.3
II L	I_I	0.046±.020	2.465±.079	−0.00±.01	80−106	+0.17	93.1
III L	I_I	0.103±.052	3.153±.233	−26.4±∞	69−2630	+0.16	95.2
IV L	I_I	0.157±.071	3.019±.295	−0.28±.36	78−137	+0.23	92.9
V L	I_I	0.629±.166	3.550±.401	−0.70±.56	84−135	+0.52	90.4
VI L	I_I	0.010±.004	2.608±.113	−0.00±.00	76−117	+0.07	93.3
VII L	I_I	0.007±.004	2.658±.120	−0.01±.00	73−118	+0.06	93.1

be overcome in the case of the Gram-Charlier by the use of the method of Least Squares, but this would increase the amount of necessary computation to a very considerable extent.

Brown and Thompson[7] advise the use of Shepherd's correction for grouping, stating that the corrected moments are "probably still valid even without high contact, if the terminal frequencies are zero," but since high contact involves the vanishing of the first n derivatives, where n is large, this does not seem to be an advisable procedure in this case and the moments have been left uncorrected for grouping.

Table I shows the principal constants for the Pearson curves for thirteen of Urban's sets of observations.[8] The data of one O (V) were omitted for the 'heavier' since they gave negative frequencies when the differences of the judgments were taken. The determining factors in selecting the proper type of curve to fit the observations are the β's which are obtained from the

[6]Arne Fisher, *The Mathematical Theory of Probability*, 2nd ed. 1, 1923, 217.
[7]Brown and Thompson, *loc. cit.*
[8]For explanation of the symbols see W. P. Elderton, *Frequency Curves and Correlation*, 1906.

moments. These quantities are always positive. If $\beta_1 = 0$ and $\beta_2 = 3$, the Normal or Gaussian Curve is the best fitting, while if only $\beta_1 = 0$, the curve is symmetrical. The probable errors of the β's are of significance, for it will be noted that in all cases it would require more than twice the probable error of the β's to carry them over to zero, or the chance of a symmetric curve in any case is less than 4.5 to 1. The criterion k,

$$k = \frac{\beta_1(\beta_2 + 3)^2}{4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)},$$

is in all cases, except IV 'heavier,' negative, which means that a Type I Curve is the most probable. In the case of IV 'heavier,' k is greater than one and is finite, which would indicate Type VI, but in this case since the skewness, -0.30, is not large and $2\beta_2 - 3\beta_1 - 6$ is small, the curve fitted was Type III. This was done for purely practical reasons, since the data were of such a nature that in determining the maximum ordinate Gamma Functions[9] of large numbers entered into the computations if Type VI was used.

In four cases, II, IV, and VII 'heavier' and III 'lighter,' the probable errors of the criterion k approached infinity which indicated that the β's were of such a nature that they were near critical values where the types change. This is best seen from Chart I, which shows the location of the β's on about one-fourth of Pearson's Diagram.[10] The lines on this chart are obtained from the values of the β's which are typical for the various types of Pearson's curves. These curves are derived from a consideration of the point binomial $(p + q)^n$, where $p + q = 1$, which when expanded gives a series which represents the probabilities of the happenings of an event, $n_1(n-1)$ . . . . 0 times when the probability of a single event is known; thus when $p = q$ we have a symmetric function which when n is large approaches the Gaussian Curve and when $p \neq q$ gives us a skewed curve, the first case being analogous to that of tossing a coin and the second that of tossing a die or a tetotem; but this does not represent all the probabilities that arise for there is another series which gives the chances of getting r, $r - 1$, ... 0 black balls from a bag containing pn black and qn white balls when r balls are drawn.[11] From differential equations derived from these series the critical values of the relationships between the β's are obtained which differentiate the types of curves and the lines on the chart are the graphs of these relationships. The point on the chart marked G repre-

[9] E. B. Wilson, *Advanced Calculus*, 1911, 378.
[10] Karl Pearson, *Tables for Statisticians and Biometricians*, 1914, lxii ff., 66.
[11] Elderton, *loc. cit.*, 36-46: Karl Pearson, Skew variation in homogeneous material, *loc. cit.*, 1895.

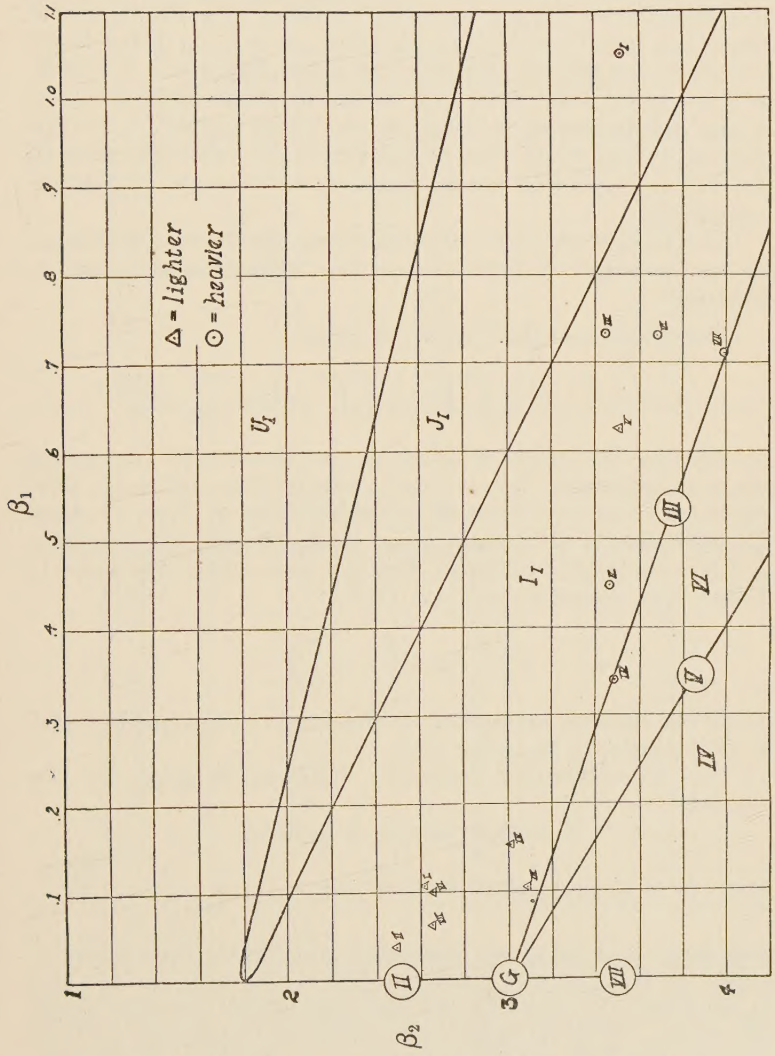

CHART I

sents the normal curve. When the point represented by the β's is very close to one of these lines the probable error of the criterion approaches infinity, *i.e.* IV 'heavier.' Quoting from a letter from Pearson in regard to this set of data, "Probably you would find it quite adequate to use a Type III curve for your data, but if you look at the chart on page 66 of the Book of Tables you will see that any one of the Types III, VI, or V, is likely to fit your data, and accordingly the probable error of the discriminating k becomes very large indeed." In this chart it will be noted that all the points are relatively close to the 'G,' *i.e.* Gaussian point, especially in the case of 'lighter' judgments.

As was to be expected all these curves are skewed, the 'heavier' in the negative direction and the 'lighter' in the positive direction.

The equation of the Type I curves is:

$$y = y_o \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a^2}\right)^{m_2}, \text{ where } \frac{m_1}{m_2} = \frac{a_1}{a_2}.$$

In this case the origin is taken at the 'mode' and the curves are not asymptotic, but cut the x-axis at $-a_1$ and $+a_2$. The curve for I 'heavier' presents a peculiar form of Type I, where the curve starts at a finite point on the x-axis, rises rapidly, and has an infinite ordinate, but the area under the curve is finite. The equation of Type II is:

$$y = y_o e^{-rx} \left(1 + \frac{x}{a}\right)^{ra}$$

where the curve is asymptotic to the x-axis at one end but cuts at a fixed point at the other end.

The Gram-Charlier frequency function is given by the equation

$$y = N[\phi_0(x) + k_3\phi_3(x) + k_4\,\phi_4(x)],$$

where N = the number of observations, $\phi_0(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$,

and $\phi_3(x)$ and $\phi_4(x)$ are the third and fourth derivatives of $\phi_0(x)$. In this case the origin is taken at the mean, and k_3 and k_4 are constants involving the third and fourth moments about the mean. These are called respectively the 'skewness' and the 'excess,' the first having a tendency to produce asymmetry, while the second tends to flatten the curve or to make it top-heavy.[12] In order to fit this curve it is necessary to have a

[12]Arne Fisher, *loc. cit.*, Parts II and III.

table giving the derivatives of $\phi_0(x)$, one of the best being that of N. R. Jørgensen in his Untersøgleser over Frequensflader og Korrelation,[13] which contains also other useful tables for curve fitting. Table II shows the values of k_3 and k_4 for the same data as the Pearson curves. It will be noted that these

TABLE II
Constants for the Gram-Charlier Curves

O	k_3	k_4
I H	+0.171	+0.020
II H	+0.118	+0.027
III H	+0.145	+0.019
IV H	+0.109	+0.023
VI H	+0.142	+0.026
VII H	+0.141	+0.041
I L	−0.056	−0.014
II L	−0.031	−0.013
III L	−0.052	+0.006
IV L	−0.062	+0.001
V L	−0.131	+0.028
VI L	−0.016	−0.016
VII L	−0.038	−0.013

TABLE III
Sum of the Squared Differences of the Observed and Theoretical Values Divided by the Theoretical

O	Normal	Pearson	Gram-Charlier
I H	60.7	24.8	65.1
II H	20.1	32.8	28.5
III H	69.9	17.7	25.6
IV H	75.7	76.3	45.5
VI H	20.2	6.1	4.0
VII H	21.9	10.9	13.3
I L	27.4	41.1	21.0
II L	9.0	25.1	6.0
III L	52.6	56.9	74.7
IV L	16.6	82.7	9.7
V L	69.8	45.0	63.5
VI L	14.9	30.6	8.6
VII L	32.5	21.1	27.0
Average	37.8	36.2	30.2

constants do not vary greatly in the various sets, an indication of homogeneity of the data.

Table III gives the values of the sum of the squared differences of the theoretical and actual values divided by the theoretical value for the Normal Curves, the Pearson Curves, and the Gram-Charlier Curves. There does not seem to be very much choice as to which type of curve gives the best fit. In

[13] N. R. Jørgensen, *Untersøgleser over Frequensflader og Korrelation*, 1916.

the case of the Normal curve, it is seen that twice it has the best fit, six times the second best and five times the poorest; while the Pearson Curves fit best five times, twice the second best, and six times the poorest. The Gram-Charlier Curves give the best fit six times, a second best fit five times, with a poorest fit only twice. If the higher moments had been adjusted by the Method of Least Squares the Gram-Charlier function would undoubtedly have given the best fit in all cases. The averages of the squared differences does not show a great amount of variation but the Gram-Charlier shows a superiority over the Pearson or the Normal forms.